PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06K 9/00	A1	(11) International Publication Number: **WO 98/02842**
		(43) International Publication Date:　22 January 1998 (22.01.98)

(21) International Application Number:　PCT/US97/06795

(22) International Filing Date:　21 April 1997 (21.04.97)

(30) Priority Data:
　08/679,085　　12 July 1996 (12.07.96)　　US

(71) Applicant: THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY OF THE NAVY [US/US]; 4555 Overlook Avenue, S.W., Code 3008.2, Washington, DC 20375-5325 (US).

(72) Inventors: PALMADESSO, Peter, J.; 15620 Rhame Drive, Montclair, VA 22026 (US). ANTONIADES, John, A.; 7522 Cherry Tree Drive, Fulton, MD 20759 (US).

(74) Agent: McDONNELL, Thomas, E.; Associate Counsel (Patents), Code 3008.2, Naval Research Laboratory, 4555 Overlook Avenue, Washington, DC 20375-5325 (US).

(81) Designated States: AU, CA, IL, JP, KR, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: INTELLIGENT HYPERSENSOR PROCESSING SYSTEM (IHPS)

(57) Abstract

The Intelligent Hypersensor Processing System (IHPS) is a system for the rapid detection of small, weak or hidden objects, substances or patterns embedded in complex backgrounds, providing fast adaptive processing for demixing and recognizing patterns or signatures in data provided by certain types of hypersensors. IHPS processes by forming a series of pattern vectors through the concatenation of the outputs of multiple sensors (10). Each sensor measures a different attribute of the system being observed, and has a consistant relationship to all other sensors. The data stream from the sensors is entered into a processing system which employs a parallel pipeline architecture. The data stream is simultaneously sent to two separate processor pipes. The first is the Demixer Module (20), which decomposes each pattern vector into a convex combination of a set of fundamental patterns. The signature pattern of a weak constituent or unresolved small target is separated from the background patterns which may hide the target. Information detailing the composition of the dimixed data patterns is sent to the Display/Output Module (40) along with information about the fundamental patterns and Filter Vectors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTELLIGENT HYPERSENSOR PROCESSING SYSTEM (IHPS)

BACKGROUND OF THE INVENTION

5   The present invention relates generally to processing signals from data provided from certain types of sensors. and more particularly to a system for the rapid detection of small. weak or hidden objects. substances. or patterns embedded in complex backgrounds. by providing a fast. adaptive processing system for demixing and recognizing patterns or

10   signatures in data provided by certain types of "hypersensors" in real time. A hypersensor is a sensor which produces as its output a high dimensional vector or matrix consisting of many separate elements. each of which is a measurement of a different attribute of the system or scene under construction. A hyperspectral imager is an example of a hypersensor. Hypersensors based on acoustic or other types of signals. or combinations

15   of different types of input signals are also possible.

      Historically there have been three types of approaches to the problems relating to the detection of small. weak or hidden objects. substances or patterns embedded in complex backgrounds.

      The first approach has been to use low dimensional sensor systems which attempt

20   to detect a clean signature of a well known target in some small. carefully chosen subset of all possible attributes. e.g.. one or a few spectral bands. These systems generally have difficulty when the target signature is heavily mixed in with other signals. so they typically can detect subpixel targets or minority chemical constituents of a mixture only under ideal conditions. if at all. The target generally must fill at least one pixel. or be dominant in

25   some other sense as in some hyperspectral bands. Also. the optimal choice of bands may vary with the observing conditions or background (e.g. weather and lighting). so such systems work best in stable. predictable environments. These systems are simpler than the high dimensional sensors (hypersensors). but they also tend to be less sensitive to subdominant targets and less adaptable.

30   The second approach has been to employ high dimensional sensor systems which seek to detect well known (prespecified) targets in complex backgrounds by using Principle Components Analysis (PCA) or similar linear methods to construct a representation of the background. Orthogonal projection methods are then used to separate the target from the background. This approach has several disadvantages. The methods used to characterize

2

the background are typically not 'real time algorithms': they are relatively slow. and must operate on the entire data set at once, and hence are better suited to post-processing than real time operation. The background characterization can get confused if the target is present in a statistically significant measure when the background is being studied. causing

5   the process to fail. Also, the appearance of the target signature may vary with the environmental conditions: this must be accounted for in advance, and it is generally very difficult to do. Finally, these PCA methods are not well suited for detecting and describing unanticipated targets. (objects or substances which have not been prespecified in detail, but which may be important) because the representation of the background constructed by these

10   methods mix the properties of the actual scene constituents in an unphysical and unpredictable way.

The more recent approach. is based on conventional convex set methods. which attempt to address the 'endmember' problem. The endmembers are a set of basis signatures from which every observed spectra in the dataset can be composed in the form of a convex

15   combination, i.e.. a weighted sum with non-negative coefficients. The non-negativity condition insures that the sum can sensibly be interpreted as a mixture of spectra. which cannot contain negative fractions of any ingredient. Thus every data vector is. to within some error tolerance. a mixture of endmembers. If the endmembers are properly constructed, they represent approximations to the signature patterns of the actual

20   constituents of the scene being observed. Orthogonal projection techniques are used to demix each data vector into its constituent endmembers. These techniques are conceptually the most powerful of the previous approaches. but current methods for implementing the convex set ideas are slow, (not real time methods) and cannot handle high dimensional pattern spaces. This last problem is a serious limitation, and renders these methods

25   unsuitable for detecting weak targets. since every constituent of a scene which is more dominant than the target must be accounted for in the endmember set, making weak target problems high dimensional. In addition, current convex set methods give priority to the constituents of the scene which are dominant in terms of frequency of occurrence, with a tendency to ignore signature patterns which are clearly above the noise but infrequent in

30   the data set. This makes them unsuitable for detecting strong but small targets unless the target patterns are fully prespecified in advance.

SUMMARY OF THE INVENTION

35

Accordingly. it is an object of this invention to provide a system for the detection of weak or hidden objects or substances embedded in complex backgrounds.

Another object of this invention is to provide a system for the accurate and fast detection of hidden objects or substances embedded in complex backgrounds via the use

5   of optical sensors.

Another object of this invention is to provide a system for the accurate and fast detection of hidden objects, substances or patterns embedded in complex backgrounds via the use of acoustic or other types of hypersensors.

Another object of this invention is to provide an efficient system for signal

10 processing capable of handling multi dimensional analysis.

Another object of this invention is to provide a set of fast algorithms for signal processing which provide a greatly reduced computational burden in comparison to existing methods.

Another object of this invention is to provide a system for processing signals from

15 hypersensors which offers true real time operation in a dynamic scenario.

A further object of this invention to provide a system for the detecting of weak or hidden objects or substances embedded in complex backgrounds which offers an adaptive learning capability.

These and additional objects of this invention are accomplished by the structures

20 and processes hereinafter described.

The Intelligent Hypersensor Processing System (IHPS) comprises sensors. a demixer module. an adaptive learning module and a means for display. IHPS forms a series of pattern vectors through the concatenation of the outputs of multiple sensors. Each sensor measures a different attribute of the system being observed, and has a consistent

25 relationship to all the other sensors. The data stream from the sensors is entered into a processing system which employs a parallel-pipeline architecture. The data stream is simultaneously sent to two separate processor pipes. The first is the Demixer Module, which decomposes each pattern vector into a convex combination of a set of fundamental patterns which are the constituents of the mixture. The decomposition is accomplished

30 using projection operations called 'Filter Vectors' generated in the second pipeline by the Adaptive Learning Module. The signature pattern of a weak constituent or an unresolved small target is separated from background patterns which may hide the target pattern in the unmixed data. A priori knowledge about the signatures of known targets can be used and approximate signatures of unknown constituents are determined automatically. Information

4

detailing the composition of the demixed data patterns is sent to the Display/Output Module along with information about the fundamental patterns and Filter Vectors.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. is a block diagram of the preferred embodiment showing the system's parallel structure.

10

Figure 2. is a representation of the data cube and the orientation of the spatial and wavelength information in X. λ and T coordinates.

Figure 3. is a logic flowchart of the operation of the prescreener

15

Figure 4. is a representation of the plane created by V_1 and V_2 during Gram-Schmidt operations.

Figure 5. is a representation of the orthogonal projections V_{1o} and V_{2o} during Gram-Schmidt operations.

20

Figure 6. is a representation of the Salient vector and plane defined by V_{1o} and V_{2o}.

Figure 7. is a representation of the 3-dimensional spanning space defined during Gram-Schmidt/Salient operations.

25

Figure 8. is a representation of the 3-dimensional spanning space showing the salient vectors.

Figure 9. is a representation of a hypertriangle convex manifold.

30

Figure 10a. is a representation of a minimized hypertriangle defined using shrink wrap method 1.

Figure 10b. is a representation of a minimized hypertriangle defined using shrink wrap method 2.

35

Figure 11. is a representation of a minimized hypertriangle defined using shrink wrap method 3.

40

Figure 12. is a logic flowchart of the operation of the adaptive learning module.

Figure 13. is a flowchart of the operation of the demixer module.

45

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring now to the Figures, wherein like reference characters indicate like elements throughout the several views. Figure 1 illustrates a block diagram of the basic architecture of the intelligent hypersensor processing system 100. A detailed description of the processing system 100 as it would operate if coupled to a hyperspectral imager (an
5   optical hypersensor) 10 is as follows. For other types of hypersensors, the spectral vectors produced by the optical system would be replaced by a vectors of other types of data elements, such as the amplitudes of different frequencies of sound. The organization of input data vectors may also vary somewhat depending on the type of sensor. Aside from these sensor-dependent variations in the type and organization of the input data, the
10   operation. capabilities. and output of the processing system would remain the same.

In a preferred embodiment the optical system is employed on an aircraft or spacecraft. As the craft flies over or in close proximity to an area of interest. Hypersensor 10 scans the scene or area of interest by taking successive snapshots of the scene below. Each snapshot constitutes a frame of spectral data. Referring now to figure 2, the spectral
15   data is scanned frame by frame 210 and displayed as variations in intensity. In the optical example, a frame 210 is the diffracted image on a 2 dimensional focal plane of a narrow slit which accepts light from a narrow linear strip of the scene. Variations of the optical sensor layout are possible.

Each frame 210 comprises multiple lines 205; each line 205 being the spectral
20   characteristic for a specific point in the scene which correlates to a specific coordinate of the area scanned. Each frame 210 is configured such that the spatial information is expressed along the X axis and wavelength information is contained in the Z direction. Figure 2 illustrates a data cube 200. created by the concatenation of successive frames 210 (different spatial strips) and represents the observed spectral data of the scene provided by
25   the hypersensor. The observed spectral data, which is used to create data cube 200 is expressed in vector form. and processed one spatial pixel. i.e. one spectrum. at a time. Each pixel is fed into a preprocessor (not shown) which performs normalization and purges bad spectral data, bad spectral data being data corrupted or otherwise useless due to incomplete spectral information.

30   Referring back to figure 1, the vector data $\{d_k\}$ is then simultaneously fed into separate processor pipes. The parallel processing architecture illustrated in figure 1 is the preferred structure. However, this system, algorithms and hardware contained herein may be employed in a system with a traditional architecture. The first processor pipe 140 comprises demixer module 20 which decomposes each vector into a convex combination

6

of a set of fundamental patterns. which are endmembers or constituents of the mixture. The decomposition is accomplished using projection operators called 'Filter Vectors' generated by the adaptive learning module 30.

5   The filter vectors $\{F_i\}$ are a set of optimal matched filters: they are the smallest vectors which will satisfy the condition:

$$F_i \cdot E_j = \delta_{ij}$$

where is δ_{ij} the Kronecker delta function (δ_{ij} equals 1 if $i = j$ and 0 otherwise) and the vectors $\{E_j\}$ are the endmembers of the scene. Thus. since each data vector (spectrum) is a convex combination of endmembers plus a noise vector N_k.

$$d_k = \sum_{j=1}^{n} c_{kj} E_j + N_k \qquad c_{kj} \geq 0$$

10

The dot product of d_k with filter vector F_J yields the coefficient of endmember E_J in the sum. plus a small error due to noise.

$$F_J \cdot d_k = \sum_{j=1}^{n} c_{kj} F_J \cdot E_j + F_J \cdot N_k \approx c_{kJ}$$

This determines the contribution which endmember J makes to the spectrum d_k. The Filter vectors are found by solving the minimization problem (smallest magnitude) described

15   above. subject to the constraint imposed by the orthogonality condition. The solution has the form

$$\underline{F} = \underline{M}^{-1} \underline{E}$$

where the filter vectors are the rows of the matrix $\underline{F}$, the endmembers (assumed to have unit magnitude) are the rows of the matrix $\underline{E}$, and the matrix $\underline{M}$ has the elements $M_{ij} =$

20    $E_i \cdot E_j$.

Filter vectors allow the signature pattern of a weak constituent or unresolved small target to be separated from background patterns which may hide the target pattern in the spectral data. Filter vectors demix the spectrum by projecting out one endmember at a time. Information detailing the composition of the demixed data spectrum is sent to

25   Display/Output module 40. along with the information about the fundamental spectra patterns. and filter vectors from adaptive learning module 30. Display/Output module 40. displays the distribution of the constituents of the scene. or transmits or stores the demixed data for later analysis.

The second processor pipe 150 comprises prescreener 50, and adaptive learning module 30. Referring now to figure 3, prescreener 50 receives data vectors from preprocessor (not shown) and generates a reduced set of vectors called the vector survivor set or survivor set 55. Survivor set 55 is then transmitted to adaptive learning module 30. Prescreener 50

5   reduces the amount of data processed by discarding spectral signatures which have been seen before 57, 52 and therefore contain no new information. This reduces the computational burden on the other elements of the learning pipeline. The vector survivor set is generally about 1000 vectors in size however the size of the survivor set can vary depending upon the conditions and applications.

10   Prescreener 50 generates the survivor set by comparing 54 the data spectra of the most recent pixel sampled with existing members of the survivor set 55. The survivor set is generated by performing dot operations in accordance to the relation:

$$d_i \bullet S_j > 1-\varepsilon$$

15   where, d_i is the newly sampled data, S_j is an existing survivor set vector, and ε is variable controlling threshold sensitivity. Here, the vectors d_i and S_j are assumed to be normalized to unit magnitude. Thus, the condition $d_i \bullet S_j = 1$ means that the two vectors are identical and the condition $d_i \bullet S_j > 1-\varepsilon$ means that they are almost identical if ε is small. Vectors

20   for d_i which meet the above condition for any survivor S_j are discarded 52 and the next vector is examined. Discarded vectors are not included in the survivor set. The value of ε, which is set by the operator or a control system, is a function of the survivor set size desired, memory length for survivor set values, desired thruput of data and the noise in the signal 56. Generally, as the value of ε is increased, the sensitivity of the system is

25   decreased. The survivor set vector, S_j, used for the above comparison is preferably chosen via the use of a structured search technique to minimize the number of dot operations while offering thorough analysis of the survivor set 55. This may be accomplished by comparing 54 the newly sampled data vector with the most recent vectors entered into the survivor set. Other search methods which minimize the number of operations necessary for

30   thorough matching are also suitable.

Figure 3 illustrates the flowchart of the logical operation of prescreener 50. The pruner 51 is a memory management device which determines when a survivor set vector should be purged from the memory. Pruner 51 monitors the survivor set 55, and adjusts the control parameters to control survivor set size. This is accomplished by setting the

value for ε 54, and the maximum allowed age of a survivor, which determines the threshold for additions to the survivor set and the time a survivor is allowed to remain in the survivor set without being regenerated.

5   Also, referring to figure 1, the survivor set data, as computed by prescreener 50 is then input 58 into adaptive learning module 30. Learning module 30, computes from the survivor set, a set of endmembers $\{E_j\}$ which together span the current scene. Endmembers are a set of fundamental patterns (vectors) from which any pattern vector in the data set can be reconstructed as a convex combination in reduced dimensional space, to within an error determined by the noise or other error criteria. The requirement that all of the

10   observed spectral vectors d_k be representable as convex combinations of conventional basis vectors, insures that the decomposition makes sense as a physical mixture of constituents, since any such mixture must have this property. The resulting patterns conform as closely as possible to signatures of actual constituents of the scene.

       Referring now to figure 12, learning module 30, employs an ordered Gram-Schmidt

15   analysis using salients to construct a reduced dimensional spanning space 125, while retaining the spectral information contained in the survivor set. The spanning space is constructed based on a spectral uniqueness hierarchy. The observed spectra of the survivor set, expressed as vector data are then projected into the spanning space 126. Computation of the endmembers is performed by Learning module 30 by projecting the survivor set

20   data into a reduced dimensional spanning space using a Gram-Schmidt/Salient analysis of the survivor set data, and employing Shrink Wrap minimization 127 to minimize the spanning space volume defined using Gram-Schmidt/Salient analysis. The endmembers are defined 128 by the vertices of the hyper triangle defined by the minimized spanning space as illustrated in figure 8.

25   **Gram-Schmidt/ Salient Analysis**

       The spanning space is defined by using a Gram-Schmidt / Salient analysis of the survivor set vectors. In the parameter vector space which contains the survivor set data, one first determines the two vectors which are furthest apart in the space, then, in the plane

30   formed by these two vectors, select two mutually orthogonal vectors which lie in the plane. These mutually orthogonal vectors are for convenience called basis vectors, for reasons made apparent below. Then select the vector in the data cube which is furthest from the plane and identify the hyperplane in which the basis vectors, and the newly selected vector, lie, and select a third basis vector such that it lies in the hyperplane and is mutually

35   orthogonal to the other two basis vectors. One repeats this process, and accumulates more

9

and more mutually orthogonal basis vectors. until the most distant remaining vector is found to be within a preselected distance of the hyperplane containing all the basis vectors. At this point. the survivor set vectors are projected onto the reduced dimensional space defined by these basis vectors. and further processing is done thereon. By reducing the

5   dimension of the vector space in which one must work, one correspondingly reduces the number of operations one must do to perform any calculation. Because none of the data vectors lie very far outside the hypervolume spanned by the basis vectors, projecting the vectors into this subspace will change their magnitude or direction very little. i.e. projection merely sheds components of each vector which were small already. Furthermore, because

10   such components are necessarily too small to correspond to significant image features, these components are disproportionately likely to be noise, and discarding them will increase the system's signal to noise ratio.

15   Gram-Schmidt \ Salient analysis of the survivor set data is performed in accordance with the following algorithm:

       a) Designate the two survivor vectors farthest apart, V_1 and V_2. Figure 4 illustrates the orientation of V_1 and V_2 and the plane that V_1 and V_2 define.
20

       b) Generate a 2 dimensional orthogonal set of basis vectors from V_1 and V_2 labeled V_{1o} and V_{2o} in the plane defined by V_1 and V_2, labeled as PV_{o12} as illustrated in Figure 5.

       c) Determine the salient vector (vector displaced farthest from plane) in relation
25   to plane PV_{o12}, defined in Figure 6 as S_1.

       d) The salient S_1 can be represented as a sum of vectors $S_{1\perp}$ and $S_{1\parallel}$, where $S_{1\perp}$ is orthogonal to the plane PV_{o12} and $S_{1\parallel}$ lies in the plane. Use the Gram-Schmidt procedure to find $S_{1\perp}$. and call this V_{3o}. V_{1o}, V_{2o} and V_{3o} now define a subspace in 3
30   dimensions. See the figure 7. representation of the subspace created by this step.

       e) Select the salient S_2 which is the survivor vector farthest from the subspace defined by step (d).

35   f) Generate a new orthogonal direction from S_2 defined as V_{4o}. V_{4o} coupled with V_{1o}, V_{2o}, and V_{3o} now defines a subspace of 4 dimensions.

       g) Steps (e) and (f) are repeated to define a spanning space of N dimensions. The distance out of the current subspace of the salient selected at each step is the maximum
40   residual error which would be incurred by projecting all of the survivors into the subspace. This decreases at each stage, until the remaining error is within a specified error tolerance. At this point the subspace construction process is complete. The value of N is the number of dimensions necessary to allow the projection of the survivor set data vectors into the subspace while at the same time preserving important but infrequent signatures.

45

10

h) Project all of the survivor set data into the spanning space defined in steps (a)-(g).

Shrink Wrap

5   Once the N-dimensional spanning space is defined using the Gram-Schmidt \ Salient analysis, a convex manifold in the form of a hypertriangle within the spanning space is generated using shrink wrap minimization. Shrink wrap minimization of the spanning space is a simple minimization operation, in which the volume of the manifold is reduced, while maintaining the condition that all of the data vectors projected into the reduced

10   dimensional space via Gram-Schmidt \ Salient operations are contained within the hypertriangle. The vertices of the hypertriangle are the endmembers, and the volume defined by the hypertriangle itself is the locus of all possible mixtures (convex combinations) of endmembers. The shrink wrap process determines good approximations of the physical constituents of the scene (endmembers), by insuring that the shape and

15   orientation of the hypertriangle conforms as closely as possible to the actual distribution of the data vectors (survivors). The survivors are assumed to be mixtures of the actual constituents. The number of endmembers is equal to the dimension of the Gram-Schmidt /Salient spanning space.

      The salients are used to guide the shrink wrap process. Referring to figure 8,

20   hypertriangle T_{S1S2S3} is defined by salient vectors, however, other vectors which include data may not be within the spanning space which T_{S1S2S3} defines, as depicted in figure 9. The shrink wrap operation must satisfy the condition that all of the data vectors projected into the spanning space must be contained within the volume defined by convex manifold T_{E1E2E3}. The shrink wrap operation starts with $T_{E1E2E3} = T_{S1S2S3}$ and then expands or

25   contracts the triangle T_{E1E2E3} by manipulating the vertices, E_1, E_2 and E_3 or by manipulating the orientation of planes that define T_{E1E2E3}, by the minimal amount to fulfill the above stated condition.

      For purposes of example, the method described above and the following methods have been found effective, however, any minimization method which maintains the

30   condition that all of the data vectors be contained within the minimized space is suitable.

      Adaptive learning module 30 generates a set of filter vectors $\{F_i\}$ and endmembers $\{E_1, E_2, E_3...E_N\}$ in accordance with one of the following procedures, or variants thereof:

Method 1

      With reference to figure 10a, find a set of endmembers $\{E_i\}$ such that each

35   endmember E_i is matched to a salient vector S_i, and is as close as possible to its salient,

11

subject to the condition that all the data vectors are inside the hypertriangle with vertices $\{E_i\}$. I.e., minimize

$$C = \sum_{i=1}^{N_D} (E_i - S_i)^2$$

subject to the constraints $F_i \cdot d_k \geq 0$ for all i and k. The filter vectors are computed from the candidate endmembers as described above. This constraint condition means that all the

5   coefficients of the decomposition of d_k into endmembers are non-negative, which is equivalent to saying that all d_k are inside $T_{E1E2...EN}$. This is a nonlinear constrained optimization problem which can be solved approximately and quickly using various iterative constrained gradient methods.

Method 2

10   Compute a set of filter vectors $\{F_{si}\}$ from the salients $\{S_i\}$, using the formulas previously provided. These vectors will not, in general satisfy the shrink wrapping constraints see figure 10b. Find a new set of Filter vectors $\{F_i\}$ such that each Filter vector F_i is matched to a salient Filter vector F_{si}, and is as close as possible to its salient filter. subject to the condition that all the data vectors are inside the hypertriangle. I.e.,

15   minimize

$$C = (F_i - F_{si})^2$$

subject to the constraints $F_i \cdot d_k \geq 0$ for all k. This is a set of independent quadratic

20   programming problems with linear constraints. which can be solved in parallel using standard methods. The decoupling of the individual filter vector calculations increases computational efficiency. Manipulation of the Filter vectors instead of the endmember vectors is equivalent to manipulating the plane faces of the triangle instead of the vertices. Given solutions for the Filter vectors. find the endmembers using the same procedure used

25   to compute Filter vectors from endmembers (the defining relationships are symmetric except for a normalization constant).

Method 3

With reference to figure 11, find an approximate centroid C_d of the set of survivor vectors. and then find the hyperplane of dimension one less than the dimension of the

30   Gram-Schmidt / Salient space which is closest to the centroid. Hyperplane 120 divides the complete subspace into two halves. and the minimization is subject to the constraint that

12

all the survivor vectors d_k must be in the same half-space as the centroid $\{c_d\}$. The normal to the optimal hyperplane 120. is F_1, the first filter vector, and the condition that all the survivors are in the same half-space is equivalent to the constraint that $F_1 \cdot d_k \geq 0$ for all

5 k. This process is equivalent to finding a vector F_1 with a fixed magnitude which minimizes the dot product $F_1 \cdot C_d$ subject to the constraint $F_1 \cdot d_k \geq 0$ for all k. As such it is amenable to solution using conventional constrained optimization methods. The hypertriangle T_{E1E2E3} can be constructed out of a set of suitably chosen optimal (locally minimal distance to the centroid) bounding hyperplanes which form the faces of the convex manifold. The normal to each face defines the associated filter vector. Again, the

10 endmembers can be determined from the Filter vectors at the end of the shrink wrapping process.

Referring to figure 12. once the endmembers and filter vectors are computed adaptive learning module 30 stores this endmember and filter vector data, along with data reflecting the survivor set, and source vectors 33 for future recall. The adaptive learning

15 module 30 then searches the survivor set for any changes. If the system detects change in the survivor set 99, the Gram-Schmidt and shrink wrap processes are repeated. This process allows the system to continually learn and adapt to changes in the environment. Endmember data and the accompanying survivor set and source data can be labeled as being consistent with a particular threat or target allowing the system to learn and

20 remember the signature of specific targets in real time 34.

Again, referring to figure 1, the filter vectors and endmember data stream are transmitted from learning module 30, to demixer module 40, for computation of the endmember coefficients. The original data set from the sensor is also transmitted to demixer module 20 through the first processor pipe.

25 Demixer module 20 may contain several processors, each of which convolves the unprocessed data vector with a different filter vector. These operations could be performed sequentially on a single fast processor, but in the best mode they are performed in parallel. The output of demixer module 20. is a vector called the endmember coefficient vector, the jth element of which indicates the fraction of the jth fundamental pattern which is present

30 in the unprocessed data vector. The endmember coefficients indicate the amplitude of the signal from the associated endmember, in a mixed spectrum.

13

Demixer module 20 convolves the unprocessed data vector and computes the endmember coefficient in accordance with the equation:

$$F_J \cdot d_k = \sum_{j=1}^{n} c_{jk} F_J \cdot E_j + F_J \cdot N_k = C_{Jk}$$

where F_j = said filter vector. d_k = said data set. c_{jk} = said endmember coefficient. N_k = noise vector and E_j = said endmember.

5   Demixer module 20 next computes the fraction coefficient 131, which tells what percentage of the photons from the given pixel are associated with the endmember in accordance to the equation:

$$c_{jk \ fraction} = \frac{c_j A(E)_j}{A(d_k)}$$

where $A(d_k)$ is the area under vector d_k, i.e. the sum of the elements of d_k.

        Figure 13 illustrates the flowchart for demixer module 20 operation including the
10   demixer module's function in the system's learning process. If the pixel information passed from the preprocessor indicates bad data or an unlearned spectrum 133, 134 demixer module 20 routes that information to the display/ output module 40 with a notation of the status of the pixel data 135, 136.

        The spectral information. filter vectors. endmember coefficients. and all is passed
15   to the display/ output module 40 for display and further processing 138. The spectral characteristics of the data is displayed in terms of endmembers and endmember coefficients in maintaining to the relation:

$$d_k = \sum_{j=1}^{n} c_{kj} E_j + N_k$$

where d_k = said data set, c_{kj} = said endmember coefficient and c is $\geq$ 0. N_k = noise and E_j = said endmember.

20   Obviously, many modifications and variations of the present invention are possible in light of the above teachings. For example this invention may be practiced without the use of a parallel processing architecture.

WO 98/02842

14

PCT/US97/06795

It is therefore understood that, within the scope of the appended claims, the invention may be practiced otherwise than as specifically described.

We Claim:

1.   A system comprising:

a sensor means for generating a plurality of data vectors;

a means for producing endmember vectors responsive to said data vectors;

wherein said means for producing comprises a module adapted to reduce the spanning

5   space of said data vectors by application of the Graham-Schmidt algorithm using salients.

2.   A sensor system comprising:

a survivor set. said survivor set comprising survivor set members;

a means for generating a data set. said data set comprising data set members wherein:

10   said means for generating compares said data set members with respective ones of said

survivor set members effective to characterize said data set members as similar or

dissimilar to said survivor set members according to a preselected criterion, and

a means for including said data set members characterized as dissimilar in said survivor

set.

15   a means for generating the endmember vectors of said survivor set:

a means for demixing said data set.

3.   An optical sensor system comprising;

      a means for collecting spectral data.

20   an adaptive learning module. said adaptive learning module computing endmember

and source spectra by performing Gram-Schmidt/Salient analysis of said spectral data said

analysis defining a reduced dimensional spanning space by:

            a)designating the two vectors contained in said data farthest apart
25   b)generating an othogonal basis set in the plane defined by said two vectors:
          c)selecting the salient relating to said plane defined by said othogonal projection:
            d)generating a new othogonal basis set based on the salient selected in step (c) said
            othogonal basis set defining a multi dimensional volume:
            continue steps (c)-(d) until the error falls within an acceptable range:

30
      said adaptive learning module further performing minimization of a convex manifold within

said multi-dimensional volume subject to the condition that all vector information from said

spectral data is contained within the minimized hypertriangle said endmembers being

defined by the vertices of said minimized hypertriangle:

16

said adaptive learning module computing said filter vectors in accordance with the equation:

$$F_i \cdot E_j = \delta_{ij}$$

5 subject to the condition that $\delta_{ij} = 1$, if $k = j$ and $\delta_{ij} = 0$ otherwise, where F_i =said filter vector and E_i = said endmember.

a demixer module, said demixer module comprising a means to compute an endmember coefficient in accordance with the equation,

$$F_k \cdot d_i = \sum c_{ij} F_k \cdot E_j + F_k \cdot N_i \quad E_j \approx c_{ik}$$

where F_k= said filter vector, d_i = said data set, c_{ij} = said endmember coefficient, N_i = 10 noise and E_i= said endmember;

a means to further process and display the spectral characteristics of said data set in terms of said endmembers and endmember coefficients said processing in accordance with the equation

$$d_i = \sum c_{ij} E_j + N_i$$

15 where d_i = said data set, c_{ij} = said endmember coefficient and is ≥ 0, N_i = noise and E_j= said endmember.

4.   An optical sensor system comprising;

a means for collecting optical spectral data,

20   a means for converting said optical data into a minimized set of data vectors comprising;

a prescreener module, said prescreener module computing a survivor set of data vectors by screening data in accordance with the equation

$$(d_i \cdot S_j > 1 - \varepsilon)$$

25 where d_i represents a new data vector and S_j represents a data vector from a survivor set;

an adaptive learning module

said adaptive learning module computing endmember and source spectra by performing Gram-Schmidt/Salient analysis of said survivor set said analysis defining a reduced dimensional spanning space by:

30

a)designating the two vectors contained in said survivor set farthest apart
b)generating an othogonal basis set in the plane defined by said two survivor set vectors;

c)selecting the salient relating to said plane defined by said othogonal basis set;

35

d)generating a new othogonal basis set based on the salient selected in step (c) said othogonal basis set defining a multi dimensional volume;
continue steps (c)-(d) until the error falls within an acceptable range,

5   said adaptive learning module further performing minimization of a convex manifold within said multi-dimensional volume subject to the condition that all vector information from said survivor set is contained within the minimized hypertriangle said endmembers being defined by the vertices of said minimized hypertriangle,

said adaptive learning module computing said filter vectors in accordance with the

10   equation:

$$F_i \bullet E_j = \delta_{ij}$$

subject to the condition that $\delta_{ij} = 1$. if $k = j$ and $\delta_{ij} = 0$ otherwise. where F_i = said filter vector and E_j = said endmember:

a demixer module:

15   said demixer module comprising a means to compute an endmember coefficient in accordance with the equation:

$$F_k \bullet d_i = \sum c_{ij} F_k \bullet E_j + F_k \bullet N_i \quad E_j \approx c_{ik}$$

where F_k = said filter vector. d_i = said data set, c_{ij} = said endmember coefficient. N_i = noise and E_j = said endmember:

20   a means to further process and display the spectral characteristics of said data set in terms of said endmembers and endmember coefficients said processing in accordance with the equation

$$d_i = \sum c_{ij} E_j + N_i$$

25   where d_i = said data set. c_{ij} = said endmember coefficient and is ≥ 0. N_i = noise and E_j = said endmember.

5.   The system of claim 2 wherein said means for generating comprises:

a prescreener module. said prescreener module characterizing said data set members as

30   similar or dissimilar to said survivor set members in accordance to the equation:

$$(d_i \bullet S_j > 1-e)$$

d_i representing said data set member and S_j representing said survivor set member and e representing a function of said system's sensitivity,

18

wherein each of said data set members which fulfill said equation is characterized as dissimilar and included in said survivor set.

6.   The system of claim 2 wherein said means for generating said endmember vectors comprises:

an adaptive learning module. said adaptive learning module computing endmember and source spectra.

7.   The system of claim 6 wherein said adaptive learning module computes endmember and source spectra by performing a Gram-Schmidt/Salient analysis of said survivor set. defining a reduced dimensional spanning space and projecting said survivor set data into said reduced dimensional spanning space by:

a)designating the two vectors contained in said survivor set farthest apart
b)generating an othogonal basis set in the plane defined by said two survivor set vectors;
c)selecting the salient relating to said plane defined by said othogonal basis set;
d)generating a new othogonal basis set based on said salient selected in step (c), said othogonal basis set defining a multi dimensional volume;
continuing steps (c)-(d) until the error falls within a desired range.

8.   The system of claim 7 wherein said adaptive learning module further defines a optimized hypertriangle, by performing a shrink wrap minimization of an objective function designated to optimize the hypertriangle in the reduced dimensional spanning space. said shrink wrap minimization fulfilling to the condition that the vector information contained in said reduced dimensional spanning space is also contained within said optimized hypertriangle.

9.   The system of claim 8 wherein said adaptive learning module computes said endmembers by defining said endmembers as the vertices of said optimized hypertriangle.

10.   The system of claim 2 wherein the means for collecting data is a hypersensor.

11.   The system of claim 2 wherein said means for generating the endmember vectors further comprises a means to compute filter vectors in accordance with the equation:

19

$$F_i \cdot E_j = \delta_{ij}$$

F_i representing said filter vectors. E_j representing said endmembers and $\delta = 1$ if $i = j$, otherwise $\delta = 0$.

12.   The system of claim 2 wherein said means for demixing comprises a demixing module. said demixing module convolving said data set and computing endmember coefficients in accordance with the equation:

$$F_j \cdot d_k = \sum_{j=1}^{n} c_{jk} F_j \cdot E_j + F_j \cdot N_k \approx c_{jk}$$

F_j representing a filter vector. d_k representing said data set. c_{jk} said endmember coefficient. N_k representing a noise vector and E_j representing said endmember.

13.   The system of claim 2 wherein said means for generating said survivor set employs a structured method which allows a minimum number of comparisons of said survivor set members to said data set information by eliminating redundant data in the data set.

14.   The system of claim 2 said means for generating further said survivor set comprises a means for managing said survivor set size as a function of age of said survivor set members.

15.   The system of claim 8 wherein said optimized hypertriangle is defined by using the filter vectors to perform shrink wrap minimization.
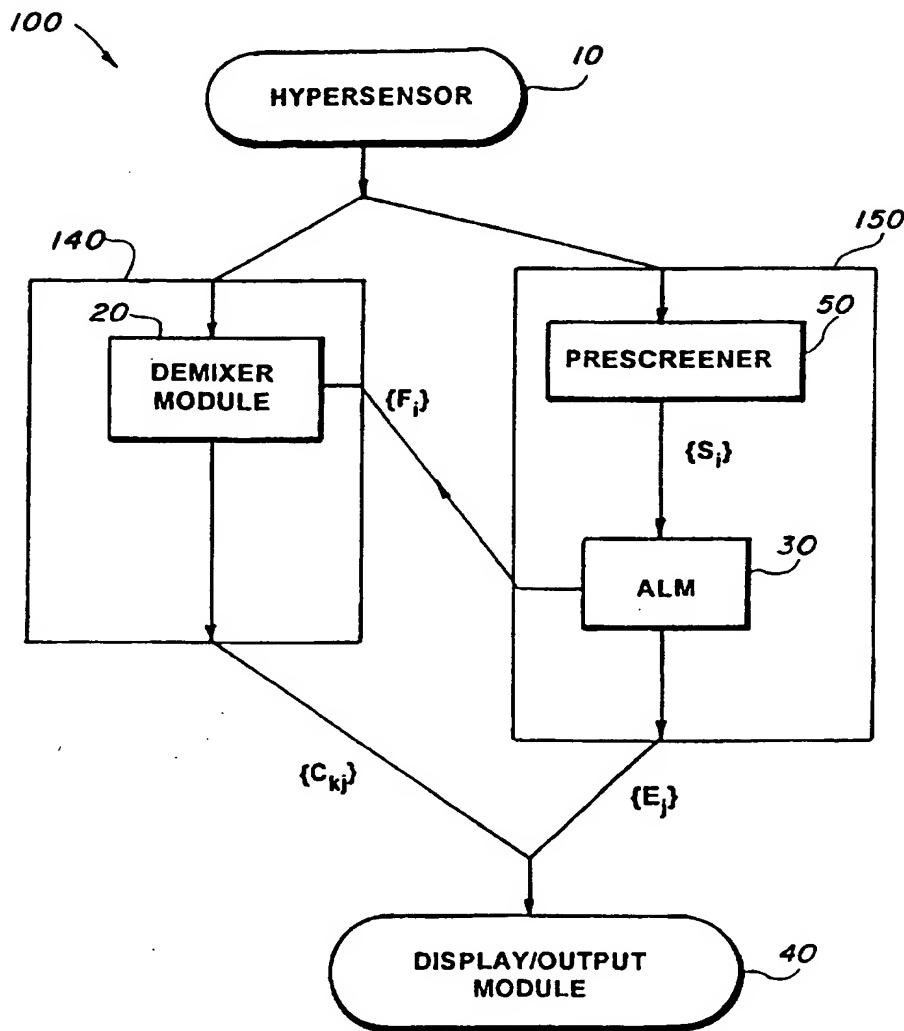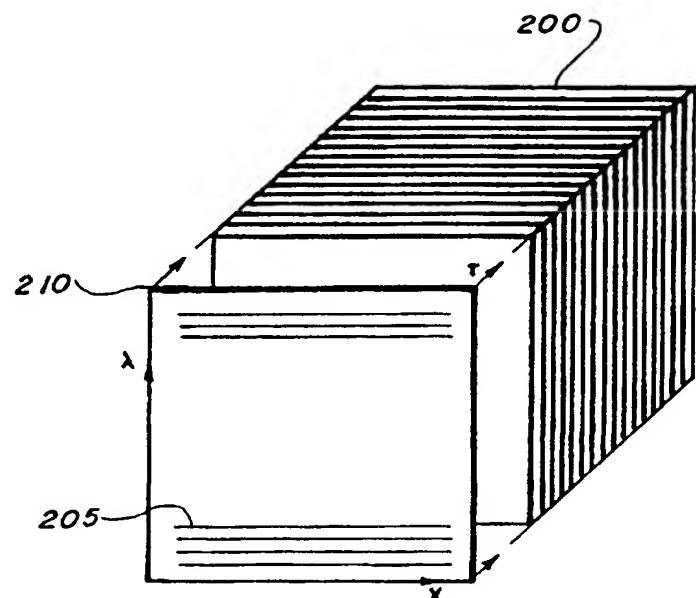
1/7

FIG . 1

FIG. 2

2 / 7

NEW DATA
$d_i \cdot S_j > 1\text{-}\epsilon$ *56*

COMPARE d_i TO
SURVIVOR SET
$d_i \cdot S_j > 1\text{-}\epsilon$ *54*

57 CONDITION
MET ?   —— YES —— DISCARD d_i *52*

NO

LABEL VECTOR
ADD d_i TO
SURVIVOR SET

SET ϵ VALUE
SET SURVIVOR SET
SIZE   *51*

SURVIVOR SET
S_j *55*

SURVIVOR
SET TO
ALM   *58*
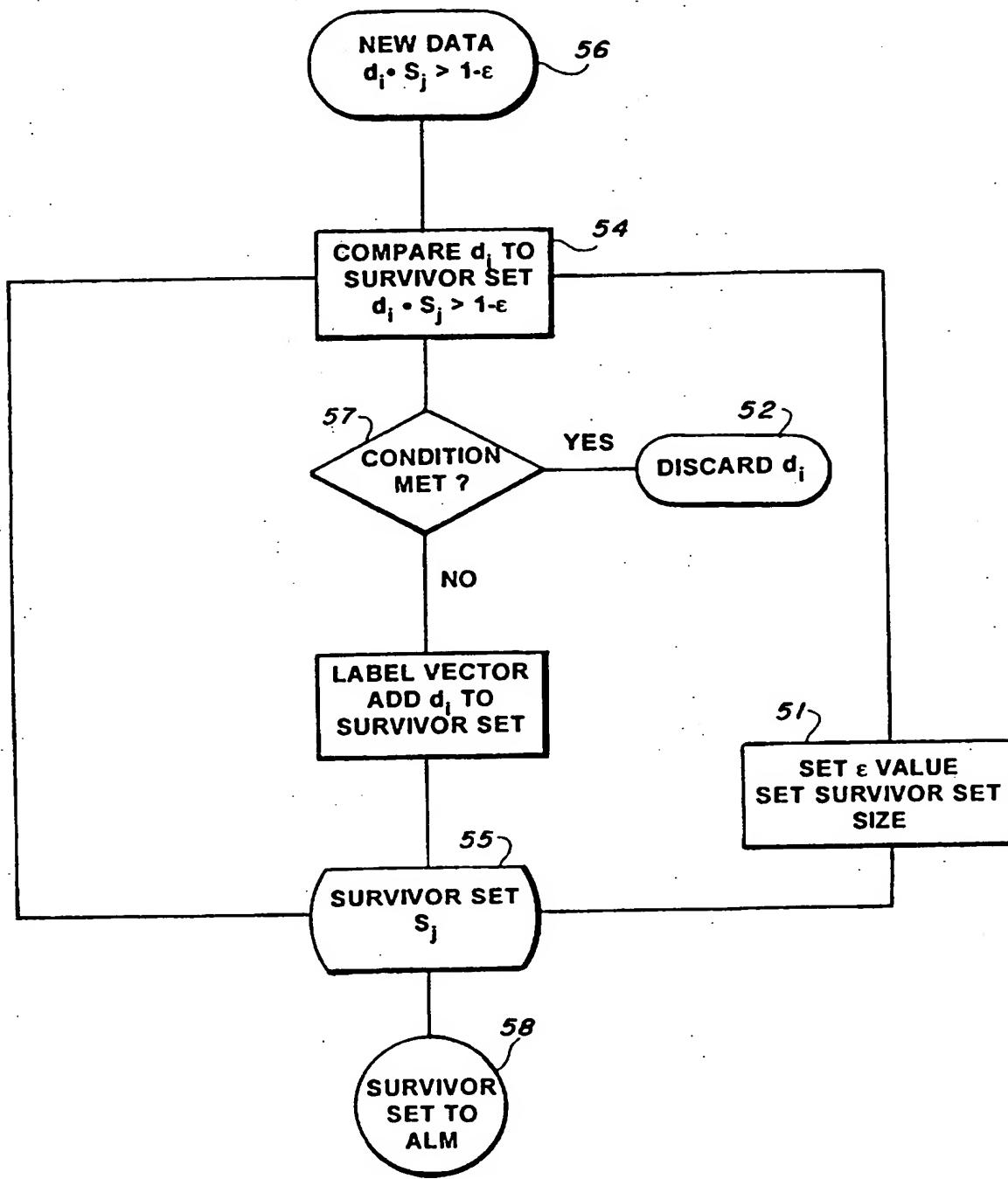
FIG. 3
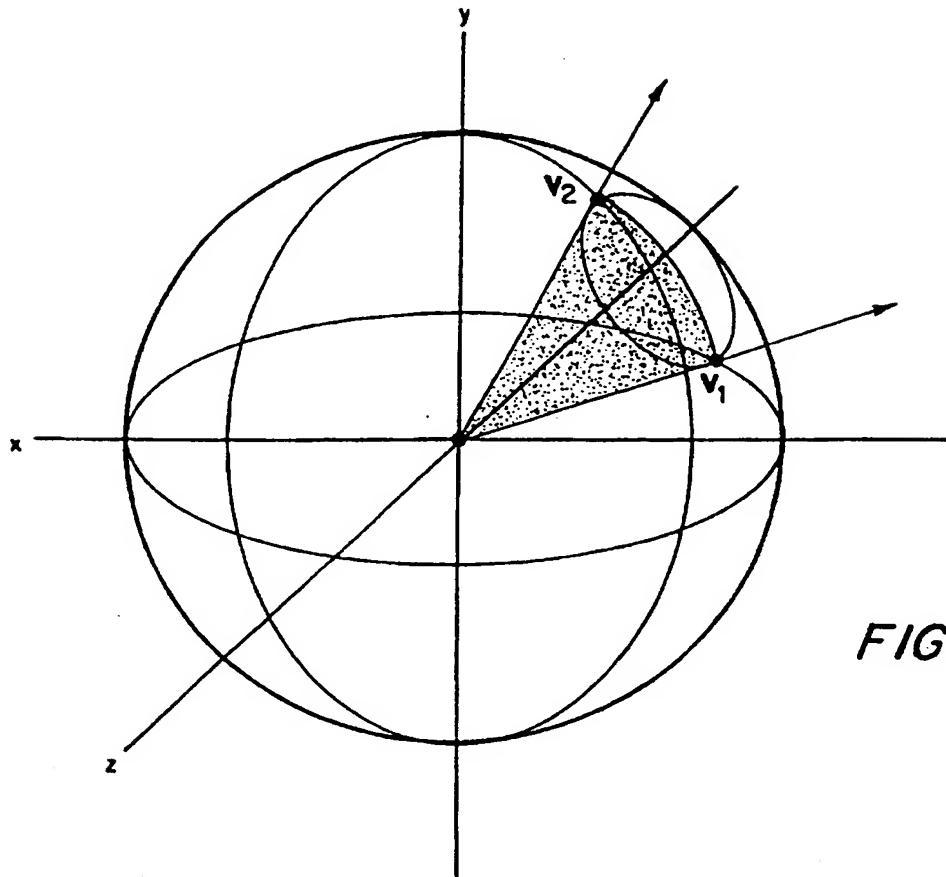
FIG. 4
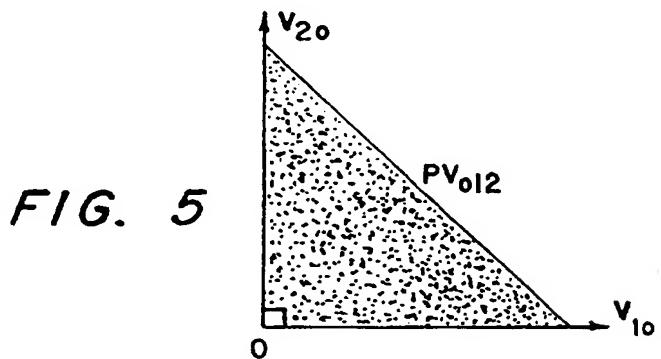
FIG. 5
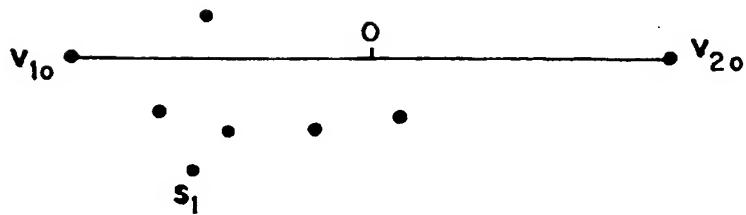
FIG. 6

4/7
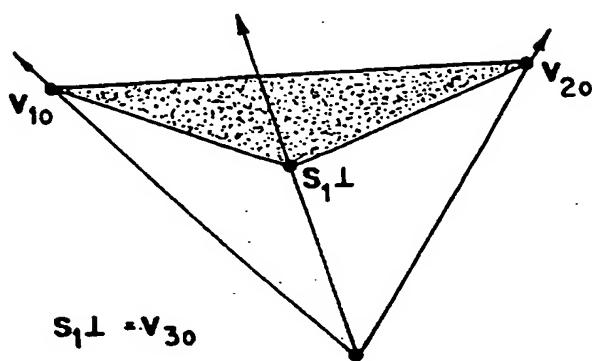
FIG. 7

$$S_1 \perp = V_{30}$$

FIG. 8

FIG. 9

$$T_{S1S2S3} = T_{E1E2E3} \, @ \, t=0$$

5/7

FIG. 10(a)

FIG. 10(b)

FIG. 11

6 / 7

30

```
          ┌─────────────────┐
          │  SURVIVOR SET   │
          │  DATA FROM      │
          │  PRESCREENER    │
          └─────────────────┘
                   │
                   ▼                          125
          ┌─────────────────────────┐
          │ PERFORM GRAM SCHMIDT/    │
          │ SALIENT TO DEFINE SPANNING│
          │ SPACE                    │
          └─────────────────────────┘
                   │
                   ▼                          126
          ┌─────────────────────────┐
          │ PROJECT SURVIVOR SET DATA INTO│
          │ GRAM SCHMIDT/SALIENT     │
          │ DEFINED BY ANALYSIS      │
          └─────────────────────────┘
                   │
                   ▼
          ┌─────────────────────────┐        127
          │ PERFORM SHRINK WRAP      │
          │ MINIMIZATION OF          │
          │ GRAM SCHMIDT/SALIENT     │
          │ SPANNING SPACE           │
          └─────────────────────────┘
                   │
                   ▼
          ┌─────────────────┐   128
          │ COMPUTE         │
          │ ENDMEMBER       │
          └─────────────────┘
                   │
                   ▼
          ┌─────────────────────────┐
          │ COMPUTE FILTER VECTORS   │
          │ MATCH ENDMEMBERS         │
          └─────────────────────────┘
```

ENDMEMBER AND
SOURCE SPECTRA
TO DISPLAY

(129)

E
ENDMEMBER
F
FILTER
VECTOR TO
DEMIXER

CHECK SURVIVOR
SET FOR NEW DATA

NEW
DATA IN
SURVIVOR
SET

YES

NO

FIG. 12
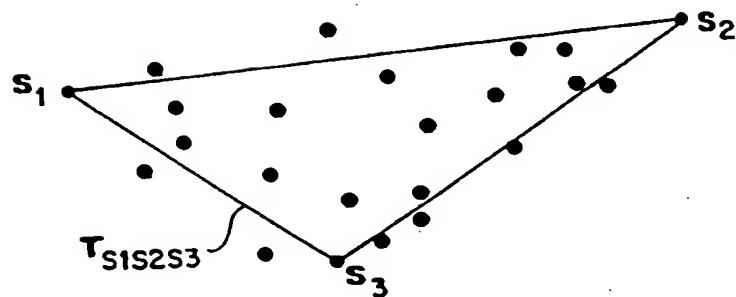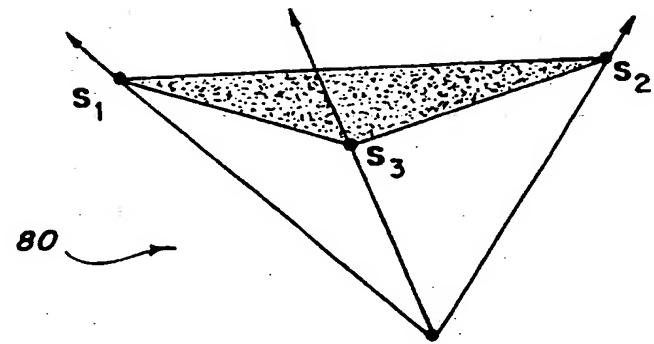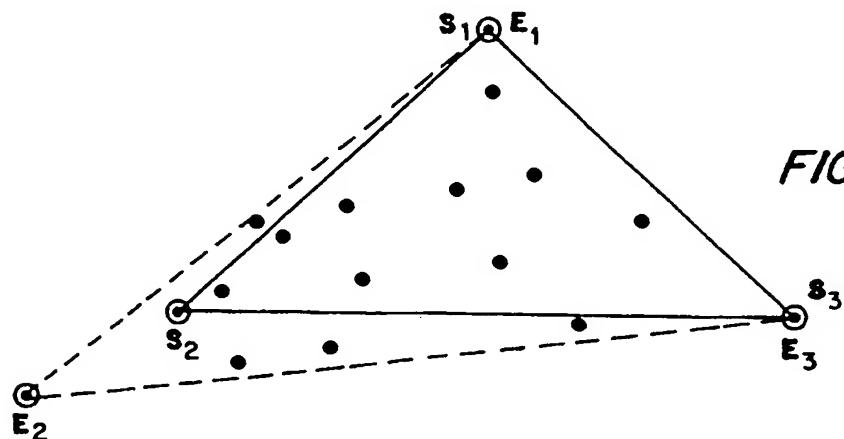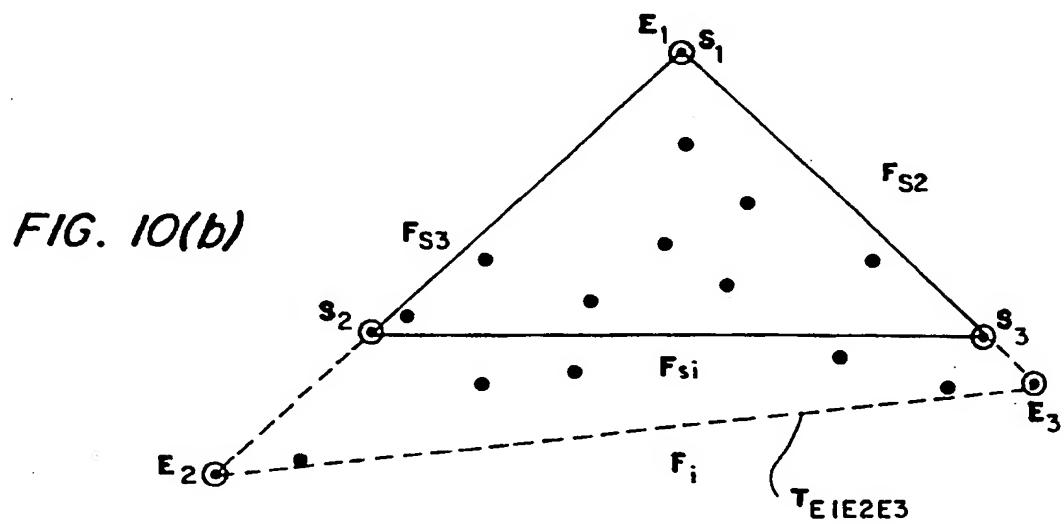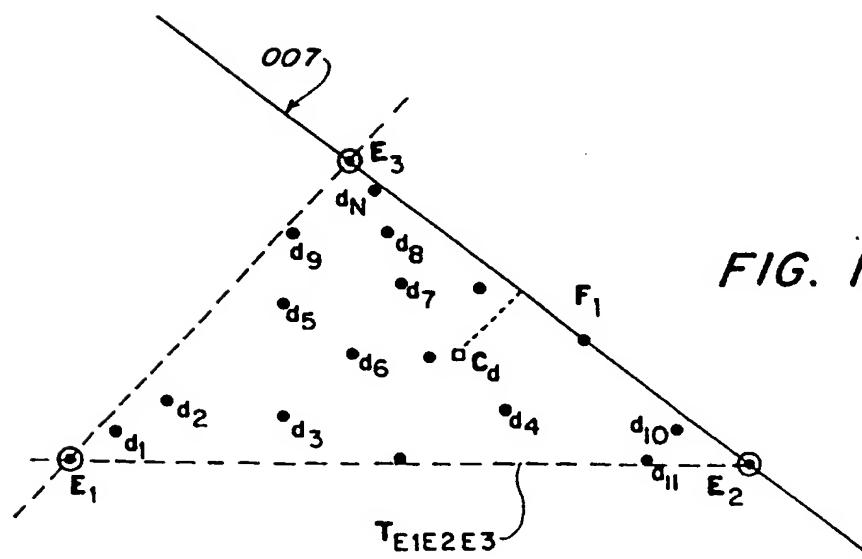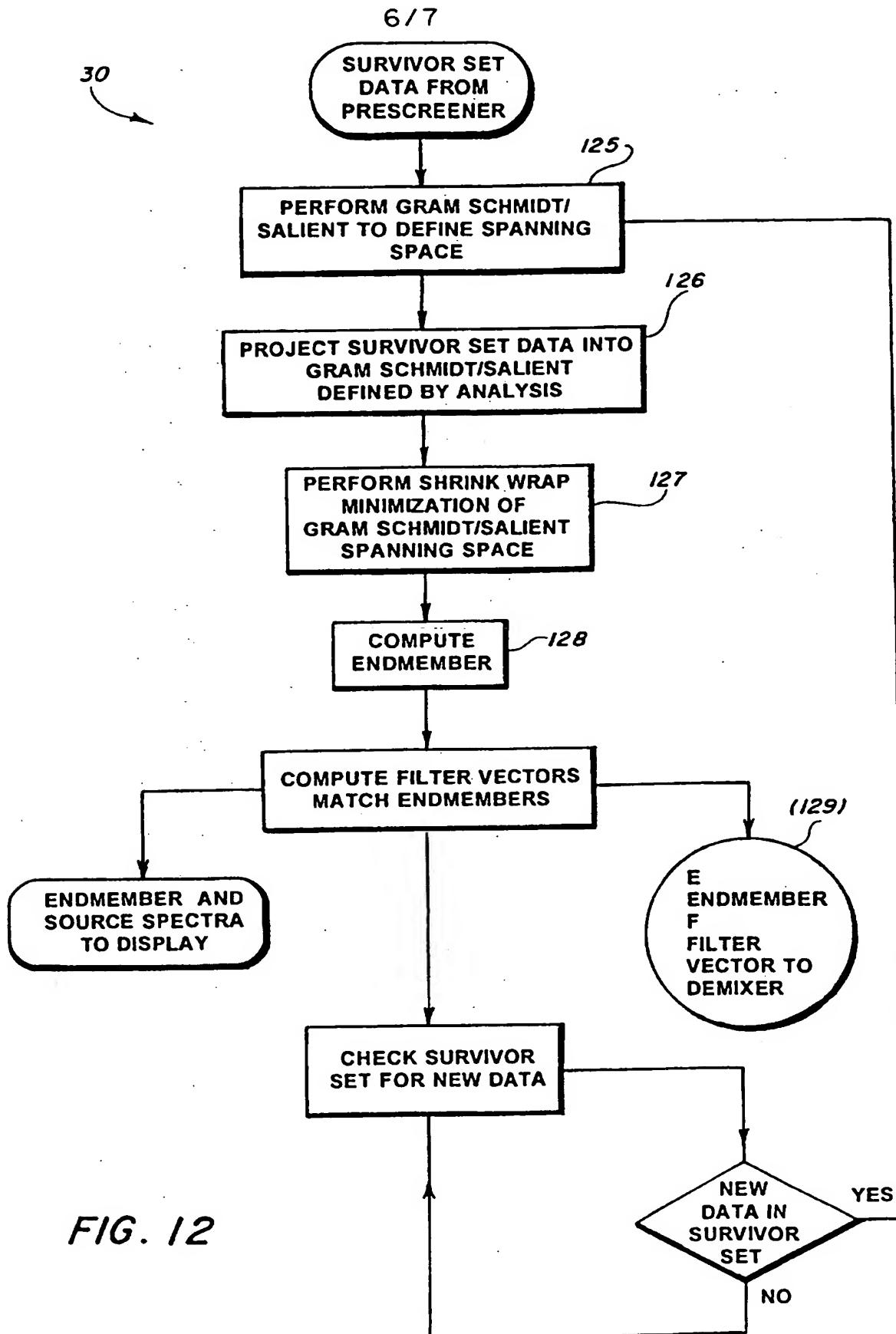
7 / 7
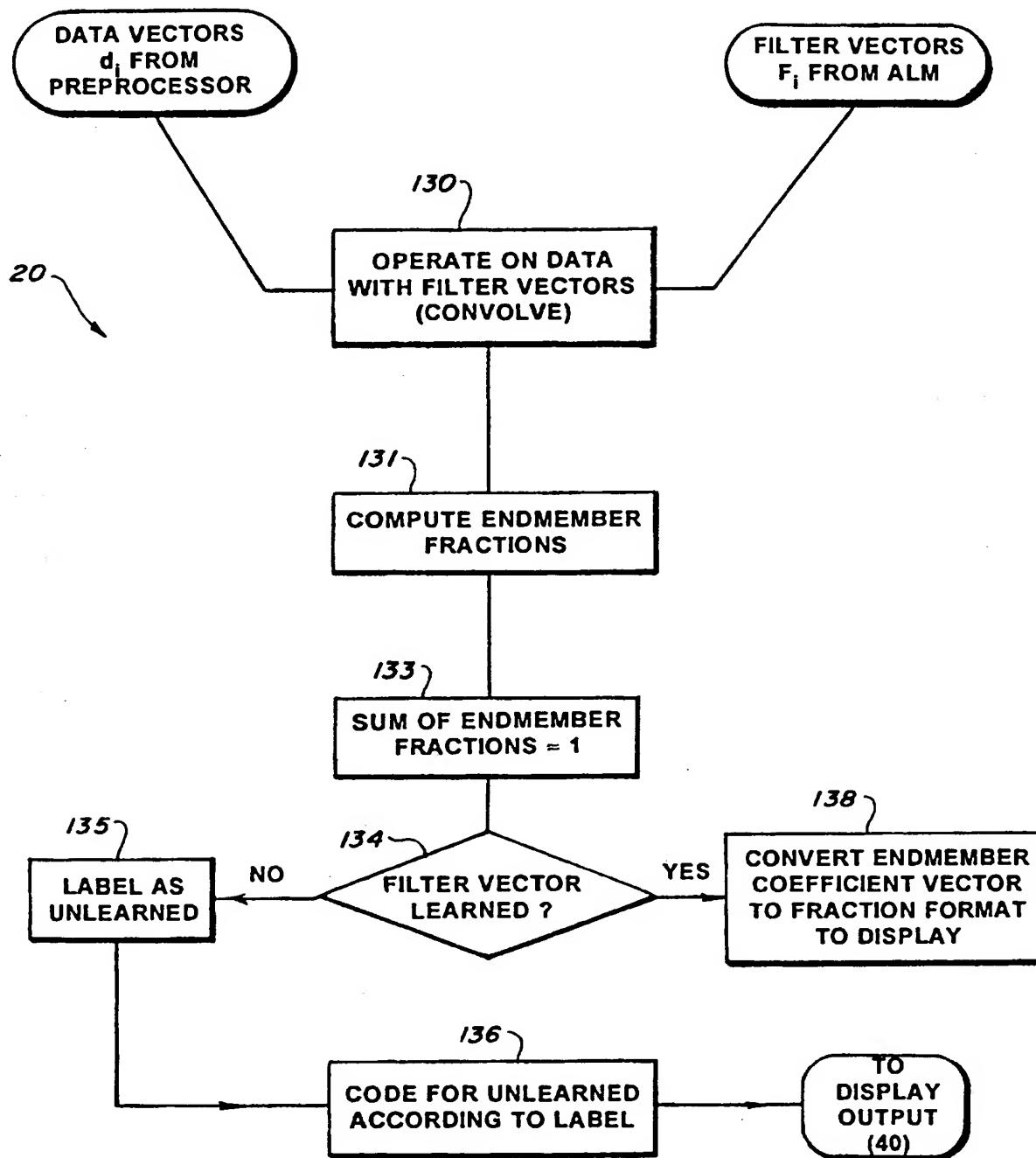
FIG. 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/06795

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6)   :G06K 9/00
US CL :382/100, 159, 249

According to International Patent Classification (IPC) or to both national classification and IPC

B.   FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   382/100, 159, 249

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SMITH, M et al. A New Approach to Quantifying Abundances of Materials in Multispectral Images IEEE Article 1994 Symposium pages 2372-2374.	1
A	CARLOTTO, M. J. Non-Linear Mixture Model and Application for Enhanced Resolution Multispectral Classification IEEE 1995 Symposium, pages 1168-1170.	1-15

☐ Further documents are listed in the continuation of Box C.   ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search

25 AUGUST 1997

Date of mailing of the international search report

3 0 OCT 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No.   (703) 305-3230

Authorized officer
CHRIS KELLEY

Telephone No.   (703) 308-0000

Form PCT/ISA/210 (second sheet)(July 1992)*

THIS PAGE BLANK (USPTO)